A Geostatistical Approach to Stochastic Frontier Models: Application to Portuguese Water Delivery Service

***Ana Sampaio*[(1)], *Carla Nunes*[(1)]**

[(1)]Dep. Matemática, Universidade de Évora

sampaio@uevora.pt, cnunes@uevora.pt

Abstract: Water is essential for all life forms. It is a fundamental resource necessary for socio-economic development and for ecological sustainability. Nowadays three of the most important problems in the human sustainable development are: the scarcity of freshwater available in natural resources, growing of water demand and the economic problem of providing safe water at a minimum cost. In the water's supply network management, cost minimization is one of the most important goals for economic authorities. In this study we use a stochastic frontier cost function (with a *single stage approach* to the estimation of inefficiency parameters) to estimate cost inefficiency measures. Geostatistical techniques are used to identify the space patterns of input variables and of the estimated cost inefficiency measure. As a result, we evaluate the possible influence of spatial patterns of inputs variables in the inefficiency measure behaviour. These approaches can be very useful for an efficient implementation of water management system. This study is applied to Continental Portuguese Water Supply System. The sample is a set of 208 observations, representing the municipalities' annual results (2000).

Key Words: *stochastic frontier cost function, spatial correlations, variogram, inefficiency, water delivery service.*

1. Introduction

Sustainable development involves a more integrated approach between environmental and economical issues and *according to the sustainable principle, all resources should be used in a manner which respects the needs of future generations* (Tietenberg, 2000).

In such an integrated system, resources allocation of a supply water service must be optimized through an alternative paradigm which ensures both ecological preservation of water reserves and social-economic goals.

Given the recent attention for the role of water service in a regulatory environment, there is a need for more empirical knowledge on the water cost structure. Recently, empirical study of water regulated utility performance has become an important policy issue in many countries. A substantial amount of research has been conducted on cost structure and efficiency analysis in the water industry (Byrnes (1985), Bosworth (1994), Cowan (1997), Cubbin and Tzanidakis (1998), Crafts (1998), Lynk (1993), Ashton (1999), Stewart (1993), Price (1993), Garcia *et al.* (2001), Bhattachatyya *et al*. (1995), Reynaud (2003) and Estache *et al*. (2002)).

As others network industries (electric power, telephone, urban transport), Portuguese water distribution is also characterized by local natural and public monopolies. Production, treatment and distribution of water in Portugal have traditionally been a public enterprise. Local authorities (municipalities) conduct, in most of the cases, the water service within each region of the country[1]: *Norte, Centro, Lisboa e Vale do Tejo, Alentejo and Algarve.*

In this study we use a stochastic frontier cost function to estimate cost inefficiency measures for municipal water distribution Portuguese service. We apply geostatistical techniques to frontier estimation methodology. In particular, we investigate if the space patterns of input variables are conditioning the space behaviour of estimated inefficiency measure.

2. Methodologies

2.1 Stochastic Frontier Approach

Originally and almost simultaneously proposed by Meuseen and van den Broeck (June, 1977), Aigner, Lovell and Schmidt (July, 1977) and Battese and Corra (1977), the stochastic frontier methodology has become increasingly popular in the analysis of productive efficiency in different sectors of economy. The essential of this approach is the existence of a composed error structure on the specification of frontier models: a symmetric component with normal distribution or the statistical noise of a traditional

[1] Madeira and Açores were not considered in this empirical study.

regression model due to data errors, omitted variables, among others, and a component with an asymmetric distribution or an unobservable random variable associated with inefficiency in production.

This paper is based on an extension of the classical stochastic cost frontier model, developed by Kumbhalkar, Ghosh, and McGuckin (1991) Reifschneider and Stevenson (1991), Huang and Liu (1994), and Battese and Coelli (1993, 1995) for the investigation of the sources of inefficiency variation. Inefficiency depends not only on the distributional specification but also on exogenous variables[2] (which are included as inputs of an inefficiency model). Following the contributions of Kumbhakar, Ghosh and McGucking (1991), and Battese and Coelli (1993, 1995) the general model is specified as[3]:

$$\ln y_i = \ln f(x_i;\beta) + \varepsilon_i$$
$$\varepsilon_i = v_i + u_i$$

where y_i denotes the output (in the case of stochastic cost function, y_i is the cost) of the i-th firm;

x_i represents a (1xK) of explanatory variables for the ith-firm;

β is an unknown vector of parameters to be estimated;

$\varepsilon_i = v_i + u_i$ [4] represents the entire error of the model: a symmetric component (V) assumed to be independent and identically distributed random errors with a normal distribution with zero mean and unknown variance σ_v^2 and an asymmetric component (U) assumed to be non-negative unobservable random variables associated with the technical inefficiency of a firm, given the levels of output and inputs. The main idea behind this methodology is that this asymmetric part of the entire error is an explicit function of a systematic component of explanatory variables (Z) and of a normal distributed unobserved random component ξ assumed to be independently distributed:

$$u_i = \delta' z_i + \xi_i$$

u_i is a realization of variables which may be independently (but not identically) distributed as non-negative truncations of the normal distribution[5] with variance given by σ_u^2, z_i is also a realization of explanatory variables associated with firm inefficiency

[2] Ownership form of production, degree of competitive pressure, input and output quality indicators, network characteristics, managerial characteristics, are examples of this kind of variables.

[3] The dual cost frontier specification is used in this paper.

[4] This is the correct signal for a cost frontier specification; for a production function the signal would be negative.

[5] Various distributions have been suggested for this term: hal-normal (Aigner, Lovell, and Schmidt, 1977), gamma (Greene,1980) and exponential (Meeusen and van de Broeck, 1977).

effects and δ is a vector of unknown parameters to be estimated. The method of maximum likelihood and the single stage approach are used to estimate simultaneously[6] all the unknown parameters (technological parameters of input variables from the systematic part of the stochastic frontier as well the inefficiency effects parameters related with exogenous variables of the inefficiency effects specification). To obtain cost efficiency estimates for each unit, we followed the methodology of Jondrow *et al.* (1982) and the predictor of the point estimator of inefficiency developed by Battese and Coelli (1988). They suggested the value of the conditional expectation of u_i as the best predictor[7]:

$$E\left(\exp\{-u_i\}\middle|\varepsilon_i\right)=\frac{1-\Phi\left[\sigma_*-\left(\tilde{\mu}_i/\sigma_*\right)\right]}{1-\Phi\left(-\tilde{\mu}_i/\sigma_*\right)}\exp\left\{-\tilde{\mu}_i+\frac{1}{2}\sigma_*^2\right\}$$

where $\Phi(.)$ is the distribution function of the standard normal random variable, $\tilde{\mu}_i$ [8] and σ_*^2 [9] are the mean and the variance of a normal-truncated conditional distribution $f(u|\varepsilon): N^+\left(\tilde{\mu}_i,\sigma_*^2\right)$.[10]

2.2 Spatial Continuity Analysis

The spatial characterization of environmental and economic variables is a fundamental step to the quantification of the environment, necessary to determine its effectiveness in achieving or increasing the sustainability of ecosystems, as well as to compare alternative plans and policies with respect to their sustainability, and to influence decision-makers.

The goal of this spatial continuity analysis is to characterize the spatial continuity of input variables and to analyse if their patterns are reproduced in the estimated

[6] The *single-stage approach* (Kumbhalkar, Ghosh, and McGuckin (1991) Reifschneider and Stevenson (1991), Huang and Liu (1994), Battese and Coelli (1993) and extended to panel data by Battese and Coelli (1995), allows the simultaneous estimation of all parameters involved (those related with the explanatory variables incorporated into the inefficiency term and the parameters associated with the regressores of the frontier cost function).

[7] This predictor corresponds to a normal-truncated normal model. The predictor corresponds to a normal-half- normal model when $\mu = 0$.

[8] With $\tilde{\mu}_i = \left(-\sigma_u^2\varepsilon_i+\mu\sigma_v^2\right)/\sigma^2$.

[9] With $\sigma_*^2 = \sigma_u^2\sigma_v^2/\sigma^2$.

[10] Following the parameterization proposed by Battese and Corra (1977), $\sigma^2=\sigma_v^2+\sigma_u^2$.

inefficiency measure.

Assuming a value $s(x_i)$ of variable S, measured at a municipality with coordinates x_i (centroid of municipality), this value can be correlated with the values measured at neighbouring municipalities during the same time period.

The variogram is a quantitative descriptive statistic that can be graphically represented in a manner which characterizes the spatial continuity of a data set. Variogram analysis consists on the definition of experimental variograms which are posterior fitted by some specific functions. The spatial continuity for a given period of time can be characterized using a spatial variogram, $\gamma_s(h)$, representing the spatial pattern (Soares 2000, Goovaerts 1997):

$$2\hat{\gamma}_{\boldsymbol{s}}(\boldsymbol{h}) = \frac{1}{Nh}\sum_{i=1}^{Nh}\left[S(x_i) - S(x_{i+h})\right]^2 \tag{1}$$

where *Nh* is the number of municipalities distance by a vector ***h***. It means that the spatial variogram is calculated by averaging one-half the difference square of the s-values over all pairs of observations with the specified separation distance and directions (vector ***h***). It is plotted as a two-dimensional graph.

3. Empirical analysis

3.1. Case study

The water distribution service data used in this study consists of a cross-section of 208 municipalities with a public service of water delivery (considering only the municipalities with owned underground or superficial water explorations), located in five Portuguese regions and surveyed in 2000. The main data source used for this study is annual statistics, made available by a State Institution - *Instituto Nacional de Estatística (INE*). Financial and physical data information on the water process is used for the estimation of the stochastic cost frontier for Portuguese municipal water delivery service. Following Stewart, M. (1993), Estache, A. and Rossi, Martín, (2002), data on operational costs (COSTS) were used to construct the dependent variable of the cost model. This variable includes total annual expenditures[11] and is defined as the sum of

[11] In 1000€.

the product of input prices and quantities for aggregate labor and capital. For the output (Q) we considered the total annual volume of water sales (in millions of cubic meters). Because estimation of a cost function requires data on input prices and as it has been difficult to obtain the prices of labour and capital, we used two proxies variables for the specification required: the average price of labour (PL), obtained dividing total wage expenses to the labour input[12] and a proxy for the price of capital (PK) obtained dividing the total expenses with capital investment to the some of the 0,6xlength of the pipe with the 0,4xnumber of well[13]. As technical variable we use the density of costumers per km^2 (DENP) calculated as the ratio between number of costumers and area (km^2). (LEN) is the size of the water system (in km), LOSS is the proportion of water loss in the distribution pipes and R_i represents the regional location indicator of the municipality (i=1,…,4). A summary of descriptive statistics concerning variables included in the general stochastic frontier cost model is listed in table 1:

Table 1

Descriptive Statistics for Variables of the Stochastic Cost Frontier Model (208 Municipalities)

Variable	Description	Mean	Std. Dev.	Min	Max
Descriptive Statistics for Variables on the General Model					
COSTS (millions of €)	Operation and maintenance costs	735750,7	708584,0	47794,8	3810716,0
Q (millions of cubic meters)	Volume of water sales	938351,0	1172449,0	125000,0	9916000,0
PL (millions of €)	Price of Labor	10725,4	10733,0	821,7	135,4
PK (millions of €)	Price of Capital	4533,3	4999,0	64,0	40811
DENP (proportion)	Proportion of costumers per km^2	139,9	314	7	2628
Descriptive Statistics for Variables Explaining Inefficiency					
LEN (Km)	Length of the pipes	190,5	204	10	1730,0
LOSS (m^3 /year)	Proportion of network losses	0,28	0,24	0	0,88

The variables are characterized by a high level of dispersion within data. Such a feature is justified to both the large standard deviations of variables and the substantial range between maximum and minimum. The great variability found on all variables reflects the municipalities' size heterogeneities.

3.2. *Stochastic Frontier Development*

[12] It is defined as total number of workers.

[13] Another process to obtain the price for the materials is proposed by Garcia and Thomas (2001): the authors used the total expenses of different inputs such as stocking, maintenance work and subcontracting divided by the distributed water volume.

Commonly, daily production of water or annual quantity of water demanded inside each municipality represents the output variable in any water delivery service. As this output is exogenous, the maximization of benefits is reach through the minimization of the cost of producing a given level of output and so the specification of a cost frontier is thus the natural choice (Estache, et al (2002).

A frontier cost function defines minimum costs given output level, input prices and a technology structure. Failure to attain the cost frontier implies the existence of inefficiency. Studies related with cost minimisation in water sector and with the investigation of sources of inefficiency include analyses by (Bhattacharyya, et al. 1995), Estache and Rossi (2002), Lynk (1993), Crampes et al. (1997).

For the estimation of cost-inefficiency of public water service of Portuguese municipalities a general stochastic cost frontier function is required[14] and expressed by:

$$\ln C_i = \ln C(y_i, w_i, x_i) + \varepsilon_i \qquad i = 1,2......N.$$
$$\varepsilon_i = v_i + u_i \qquad \mathrm{u_i} \geq 0;$$

Where C_i denotes the total cost for the i municipality (i=1,…N), $y \in R_+^1$ represents a single output or the quantity (in cubic meters) of water delivered, $W \in R_{++}^2$ denotes the prices of inputs L and K where L means labour and K means the capital, $X \in R_+^P$ represent technical variables as each municipality faces different environments[15]and ε_i =v_i+u_i is a composed and stochastic error term. For the study of the factors determinants of technical inefficiency, it is considered a vector $Z \in R_+^M$ of three exogenous variables: a network variable, an environmental variable and a geographical variable.

The unknown cost frontier is approximated by the following Cobb-Douglas specification[16]:

$$\ln\left(COSTS_i\right) = \beta_0 + \beta_1 \ln Q_{1,i} + \beta_2 \ln PL_{2,i} + \beta_3 \ln PK_{3,i} + \beta_4 \ln DENP_{4,i} +$$
$$+\varepsilon_i$$
$$i = 1,....N = 208.$$
$$\varepsilon_i = v_i + u_i$$

[14] Following the properties of a cost function (concave and linearly homogenous in input prices and non-decreasing in input prices and output).

[15] In this study it is only considered the density of population as technical variable.

[16] The Cobb-Douglas functional form, in comparison to the translog form, reduces substantially the number of explanatory variables; the major limit is that the value of the economies of scale is constant.

Where i indicates a Portuguese municipality and the β_i are unknown parameters to be estimated. The inefficiency effects model is expressed by:

$$u_i = \delta_0 + \delta_1 Len + \delta_2 Loss + \delta_3 R_1 + \delta_4 R_2 + \delta_5 R_3 + \delta_6 R_4 + w_i$$

For the maximum-likelihood estimation of the unknown parameters, we adopted the computer program FRONTIER 4.1 (Coelli, 1996)[17]. The maximum likelihood estimates of the parameters and the t-ratios are presented in table 2:

Table 2

Maximum Likelihood Estimates of Cost Frontier Model

Parameter and Variable	**Model I (no inefficiency effects)[18]**		**Model II (normal-truncated error component)**	
	Coefficient	**t-ratio**	**Coefficient**	**t-ratio**
B_0 intercept	0,71	0.9468	2,24	3,30
B_1 lnQ	0,49	10,70	0,25	5,43
B_2 lnPL	0,30	5,19	0,20	4,13
B_3 lnPK	0,32	9,03	0,45	13,49
B_4 lnDENP	0,11	3,09	0,10	3,05
σ^2			0,14	8,97
γ[19]		0,9999		38329,35
δ_0 intercept			-0,50	-1,65
δ_1 lnLEN			0,49	10,57
δ_2 Ln(LOSS+1)			0,42	2,05
δ_3 R1			-0,41	-2,76
δ_4 R2			-0,27	-1,73
δ_5 R3			-0,37	-2,53
δ_6 R4			-0,36	-2,38
LLF	-159,66		-90,73	

Almost of the ML estimates of Model II for the coefficients associated with input prices, output and technical variables, are significantly different from zero at the five

[17] This program was employed to simultaneously estimate the parameters of the stochastic cost frontier and the technical inefficiency effects model.

[18] These estimates are obtained in a way that all the values are unbiased estimates of the coefficients β and σ^2 except the value for the intercept which is biased because of the non zero expectation of u_i; they are used as starting values in the interactive process to obtain the ML estimates for generalized truncated-normal model.

[19] Where $\gamma = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_v^2}$. The parameter γ lies between 0 and 1. It provides good starting values for the iterative maximization routine which is used to calculate the maximum likelihood parameter estimates. When $\gamma = 0$, all the deviations from the frontier are due to statistical noise; when $\gamma = 1$ all deviations are due to inefficiency. The null hypothesis that inefficiency effects are not random is expressed by $H_0 : \gamma = 0$ and further, the null hypothesis that the inefficiency effects are not influenced by the level of the explanatory variables is expressed by $H_0 : \delta = 0$.

percent level. These estimates are asymptotically efficient. Most of the β estimates and δ estimates are statistically significant in both models. The *ML* estimate for γ is almost 1 (0,9999999), which indicates that the majority of residual variation is due to the inefficiency effects or that the random error is near to zero and that the stochastic frontier is not significantly different from the deterministic frontier. Restrictions associated with these two models are assumed and the adequacy of these models is tested using likelihood-ratio tests[20]. The results concerning the choice of the preferred model are summarized in table 3:

Table 3

Choice of the Preferred Model

Restrictions	Model Description	Log likelihood	λ	Critical Value	Decision
1. None	Model II- (Normal-Truncated)[21]	-90,73			
2. H₀ $\gamma = \delta_0 = \delta_1 = ...\delta_6 = 0$	Model I	-159,66	137	$\chi^2_{8;\alpha=0,05} = 15,5$	Rejected

These results indicate that the traditional average response function (Model I) is not an adequated representation of the data. The null hypothesis that the inefficiency effects are absent (Model I) or network characteristics, environmental factors and geographical location do not influence the municipality's technical inefficiency, is strongly rejected at the 5% level of significance in favour of the preferred Model II.

The positive signs on the estimate of the coefficients of capital price and labour price were as expected; the estimated coefficient associated with the length of the pipe (δ_1) has a non expected positive sign (It was expected that efficiency would increase with the length of the pipe, rather than decrease, because of scale economies) and the positive sign associated with population density in the area served was also not expected.

[20] All relevant hypotheses were tested using the generalized likelihood-ratio statistic $\lambda = -2\{\ln L(H_0) - \ln L(H_1)\}$, where Ln($H_0$) and Ln ($H_1$) are the values of the log-function under the null and alternative hypotheses, respectively. This statistic has a mixed chi-square distribution, with degrees of freedom equal to the number of restrictions imposed under the null hypothesis.

[21] We had tested the null hypothesis (H_0: μ=0) that the simpler half-normal model was a good representation of the data, given the truncated-normal model through a generalized-ratio test. We adopted the generalized truncated-normal model, although the test statistic was significant and the null hypothesis of μ=0 accepted (LL= -91,23 , λ=1 and the critical value equal to 1).

3.3. Spatial Continuity Analyses

For a spatial study it is necessary to have well defined point coordinates. A municipality is not a specific point, but an area. In this study it is considered as a geographic coordinates of an area the centroid coordinates.

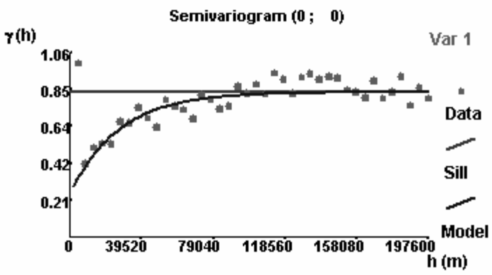
Figure 1 shows the centroids' municipalities locations where the input variables were collected.

Figure 1. Centroids' municipalities locations.

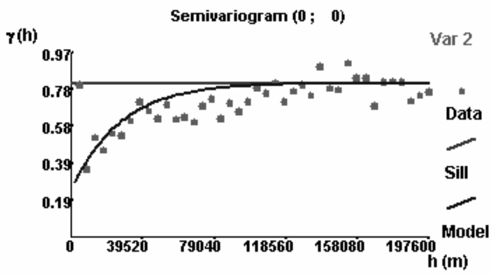
Note that not all the municipalities are represented, because we only considered those with owned water capture systems. Also in north of Portugal we have more municipalities with small areas, justifying the high density of black points in this north area, comparing with south of the country (with extensive municipalities areas).

The descriptive analyses were done in the previous section, and here, the goal, is the spatial patterns characterizations, the definitions of theoretical parameters of the fitted model and their interpretations.

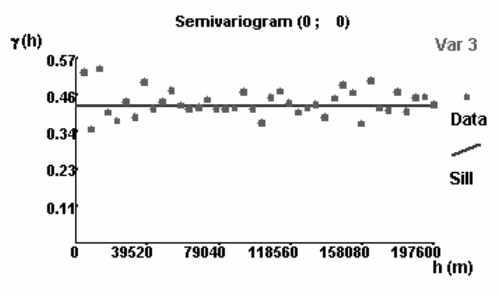
Figure 2 presents the experimental variograms of all input variables considered in the stochastic frontier model and, also, the experimental variogram of the estimated cost inefficiency measure.

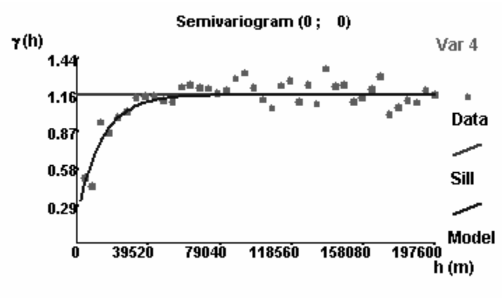


a) ln(COSTS)



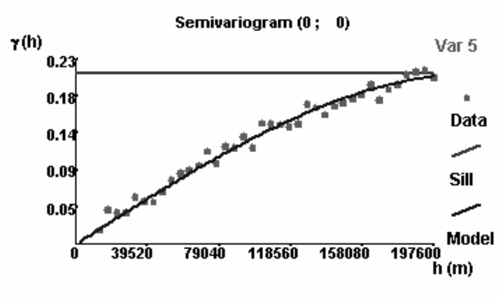
ln(Q)

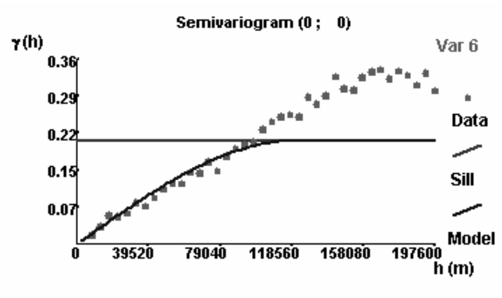




c) ln(PL)

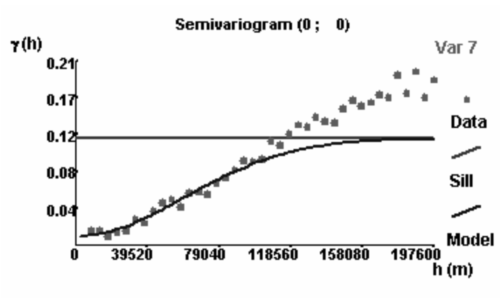
d) ln(PK)

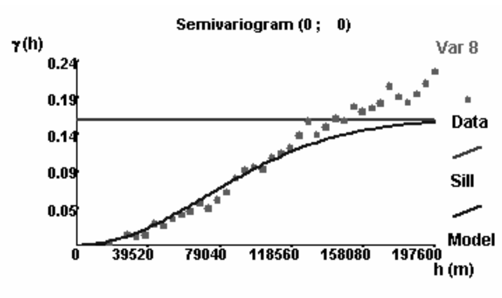




e) R1

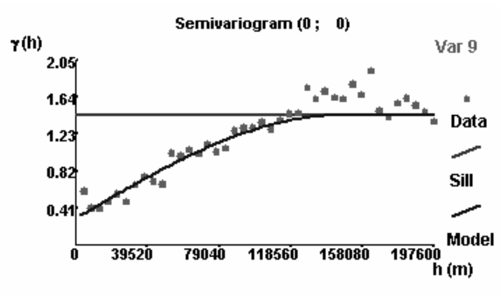
f) R2

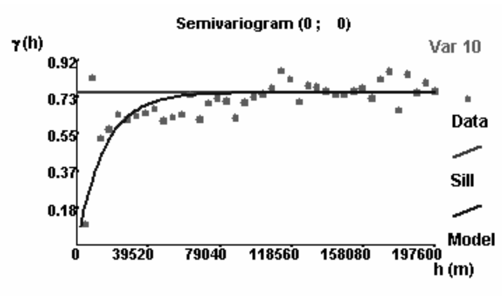




g) R3

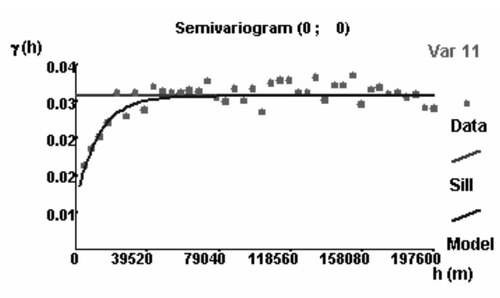
h) R4

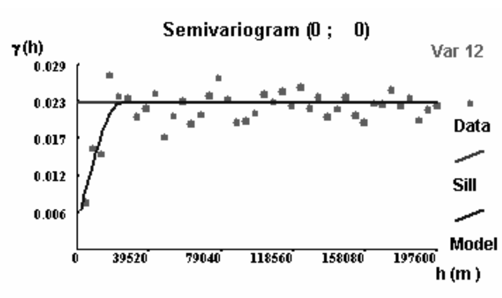




i) ln(DENP)

j) ln(LEN)





k) ln(LOSS+1)

l) Ef

Figure 2. Experimental variograms

For each variable, a) to i), the back points are the experimental variograms, the horizontal lines are the sill of variograms (representing the total semi-variance of variable), and the adjusted lines are the theoretic models fitted to each experimental variogram. All variables present some spatial structures with different dimensions of neighbourhood correlations. Only variable ln(pl) does not present any spatial structure (figure 2.c). Only omnidirectional variograms were successfully computed.

The fitted models parameters are presented in table 4.

Table 4

Fitted Models Parameters

	Nugget Effect-first structure	C1- Second structure contribution	Total Semi-Variance	Range	Model
Ln(COSTS)	0.25	0.582	0.832	80 000	Exponential
Ln(Q)	0.25	0.563	0.813	80 000	Exponential
Ln(PL)	Pure nugget effect- No spatial correlation				
Ln(PK)	0.2	0.957	1.157	40 000	Exponential
R1	0	0.213	0.213	220 000	Spherical.
R2	0	0.203	0.203	120 000	Spherical.
R3	0.1	0.11	0.12	145 000	Gaussian
R4	0	0.161	0.161	175 000	Gaussian
Ln(DENP)	0.3	1.139	1.439	145 000	Spherical.
LN(LEN)	0	0.755	0.755	45 000	Exponential
Ln(LOSS+1))	0.01	0.022	0.032	37 000	Exponential
EF	0.004	0.019	0.023	24 000	Spherical

The nugget effect represents the non modelling part of the spatial phenomena: it can be due to sampling errors, not appropriated sampling scale, analytical errors, among others. If the experimental variogram appears to have a non-zero intercept on the vertical axes, then the model may need a nugget effect component. It can be interpreted as a proportion of the total semi-variance. For instance, ln(COSTS) has a nugget effect of 30% of total semi-variance (0.25/0.832) of non-modelled part of variance. The correspondent variogram only justify 70% of variable dispersion behaviour (with an aureole correlation of 80 000 meters- range- and using an exponential model).

In this study we can observed that we have variables with 0% of nugget effect (all the behaviour dispersion can be spatial justified), and also, as already mention, with high values of nugget effects. The range in the input variables varies between 0 (pure nugget effect- ln(PL)) and 220 000 meters, which suggest an high level of heterogeneity between aureole correlation dimensions (range). The type of fitted models also varies for each variable. It means that, in spatial terms, the input variables have some spatial patterns, although not always similar.

The estimated inefficiency measure present the lower estimated range (not considering the ln(PL)variable), with a nugget effect of 17%, and using a spherical model.

4. Conclusions

The purpose of this research was to study the effects of exogenous variables upon the municipals' technical efficiency cost levels of Portuguese public water distribution service using a Cobb-Douglas stochastic cost frontier function.

The technical inefficiency component has been assumed to follow a positive truncated normal distribution and statistical tests were carried out to justify the choose of the technical inefficiency effects model.

It was done a spatial characterization of input variables and of the inefficiency measure, concluding that they have different spatial patterns.

This analysis only characterizes this case-study with these input variables. We can not justify the general influence of the spatial patterns of input variables in the estimated stochastic frontier results. It represents a spatial picture of this specific model, not pretend to extent to more general situations.

Further Development are request in order to simulate several hypothetical scenarios to try to establish general conclusions (using random spatial inputs variables, different stochastic frontiers specifications, different functional forms, among others,…).

References

Aigner, D.J., Lovell, C.A.K. and Schmidt, P. (1977). "Formulation and estimation of stochastic frontier production function models", *Journal of Econometrics* **6**, pp. 21-37.

Ashton, J. (1999). "Economies of scale, economies of capital utilization and capital utilization in the English and Welsh water company". School of Finance and Law, Bournemouth University, *Working Papers Series*, nº17.

Battese, G. E., and S.S. Corra (1997). "Functional Forms of Stochastic Frontier Production Functions and Models for Technical Inefficiency Effects: A Comparative

Study for Wheat Farmers in Pakistan", *Journal of Productivity Analysis*, **8**:4 (November), 395-414.

Battese, G. E., and T.J. Coelli. (1992). "Frontier production Functions, Technical Efficiency and Panel Data: With Application to Paddy Farmers in India", *Journal of Productivity Analysis*, **3**, 153-169.

Battese, G. E., and T.J. Coelli. (1995). "A model for technical inefficiency effects in a stochastic frontier production function for panel data". *Empirical Economics*, **20**, 325-332.

Battese, G.E., and G.S. Corra (1977). "Estimation of a Production Frontier Model: With Application to the Patoral Zone of Eastern Australia". *Australian Journal of Agricultural Economics*, **21**, 169-79.

Bhattacharyya A., T. R. Harris, R. Naraymanam, e K. Raffiee. (1995). "Specification and Estimation of the Effect of ownership on the Economic Efficiency of the Water Utilities. *Regional Science and Urban Economics* **25**, 759-784.

Bosworth, D.,P. S. and J.Roe, 1994. "Water and Sewerage industries: general efficiency and the potential for improvement". *OFWAT.*

Byrnes, Patricia (1985). "Ownership and efficiency in the water supply industry: an application of the nonparametric programming approach to efficiency measurement". PhD Doctoral Thesis.

Cowan, S. (1997) Competition in water industry. *Oxford Review of Economic Policy*, vol. **13** (1):83-92.

Crafts, N., 1998. "Water and Sewerage industries general efficiency and potential for improvement". *OFWAT.*

Crampes, C., Diette, N., and Estache, 1997. “What Could regulators learn from yard stick competition ?”Lessons for Brazil’s water and sanitation sector. Mimeo, the *World Bank*.

Cubbin and Tzanidakis (1998). Regression versus data envelopment analysis for efficiency measurement:\an\ applicaztion to the England and Wales regulated water industry. *Utilities Economic* **7**, 75-85.

Estache, A. ans Rossi, M. (2002). “How Different is the Efficiency of Public and Private Water Companies in Asia?”. *The World Bank Economic Review*, vol.**16**, nº1, 139-148.

Garcia, Serge e Thomas Alban. (2001). “The Strutucture of Municipal Water Supply Costs: Application to a Panel of French Local Communities”. *Journal of Productivity Analysis*, **16**, 5-29.

Goovaerts, P. (1997). “Geostatistics for natural resources characterization”. *Oxford University Press*, New York.

Greene, W. (1980). “Maximun Likelihood Estimation of Econometric Frontier”, *Journal of Econometrics*, 13, pp. 27-56.

Huang, J.T.Liu, (1994). “Estimation of a non-neutral Stochastic Frontier Production Function”. *Journal of Productivity Analysis*, 5, 171-180.

Kumbhakar, S.C.,S.Ghosh, and J.T.McGuckin (1991). “A Generalized Production Frontier Approach for Estimating Determinants of Inefficiency in U.S. Dairy Farms”. *Journal of Business and Economic Statistics*, **9**:3, 279-86.

Lynk, E.L., (1993). “Privatization, joint production and comparative efficiencies of private and public ownership: The UK water industry”. *Fiscal Studies*, vol.**14**, nº2, pp 98-116.

Meeusen, W., and J.van den Broeck (1977). “Efficiency Estimation from Cobb-Douglas Production Functions with Composed Error”. *International Economic Review*, **18**:2, 435-44.

Reifschneider D., Stevenson R., 1991. “Systematic Departures from the Frontier :A Framework for the Analysis of Firm Inefficiency”. *International Economic Review*, **32**:715-723.

Reynauld, A., (2003). “An Econometric Estimation of Industrial Water Demand in France”. *Environmental and Resource Economics*, **25**:213-232.

Soares A. (2000). “Geoestatística para as ciências da terra e do ambiente”. IST *Press Lisboa*, 206 p.

Stewart, M., (1993). “Modelling Water Costs 1992-1993:Furhter Research into the Impact of Operating Conditions on Company Costs”. O*FWAT Research Paper nº* ***2****, Office of Water Services*, Birmingham.

Tietenberg, T., 2000. “Environmental and Natural Resource Economics”. *Addison-Wesley Longman, Inc*.